

Umgang mit Forschungsdaten

Checkliste für Antragstellende zur Planung und zur Beschreibung des Umgangs mit Forschungsdaten in Forschungsvorhaben

Anmerkung

Die DFG stellt ausführliche [Informationen zum Umgang mit Forschungsdaten in DFG-Projekten](#) bereit. Insbesondere muss bei der Antragstellung die hier vorliegende [Checkliste](#) beachtet werden.

Die Checkliste wurde mit Kommentaren und Erläuterungen (blaue Textteile) vom Zentrum für nachhaltiges Forschungsdatenmanagement der Universität Hamburg ergänzt. Bei Fragen sprechen Sie uns gern an. www.fdm.uni-hamburg.de (April 2022)

Diese Checkliste unterstützt Sie, wesentliche Aspekte des Umgangs mit Forschungsdaten strukturiert zu beschreiben sowie die für die Umsetzung benötigten Ressourcen und Kompetenzen sichtbar zu machen. Bitte nehmen Sie zu den unten genannten Themenfeldern im Antrag unter Punkt 2.4 Stellung.

Zu Forschungsdaten zählen u. a. Messdaten, Laborwerte, audiovisuelle Informationen, Texte, Surveydaten oder Beobachtungsdaten, methodische Testverfahren sowie Fragebögen. Korpora, Software und Simulationen können ebenfalls zentrale Ergebnisse wissenschaftlicher Forschung darstellen und werden daher ebenfalls unter den Begriff Forschungsdaten gefasst. Da Forschungsdaten in einigen Fachbereichen auf der Analyse von Objekten basieren (z. B. Gewebe-, Material-, Gesteins-, Wasser- und Bodenproben, Prüfkörper, Installationen, Artefakte und Kunstgegenstände), muss der Umgang mit diesen ebenso sorgfältig sein und eine fachlich adäquate Nachnutzungsmöglichkeit, wann immer sinnvoll und möglich, mitgedacht werden. Steht die Nachnutzbarkeit der entstehenden Forschungsdaten in engem Zusammenhang mit Objekten, so bitten wir Sie, auch entsprechende Angaben zu diesen zu ergänzen.

Bitte berücksichtigen Sie die in Ihrer Fachdisziplin existierenden Standards und ggf. bestehende fachspezifische Empfehlungen und Angebote existierender Infrastrukturen (z. B. Datenrepositorien, Archive oder Sammlungen). Einen Überblick über existierende Strukturen bietet das Portal für Forschungsinfrastrukturen RIsources (<https://risources.dfg.de>) sowie das Verzeichnis von Forschungsdatenrepositorien re3data (<http://re3data.org>).

Weitere Informationen zum Thema und fachspezifische Empfehlungen finden Sie unter: www.dfg.de/antragstellung/forschungsdaten/

Checkliste zum Umgang mit Forschungsdaten

Die DFG erwartet in Anträgen auf Förderung die Beantwortung der Fragen aus der folgenden Checkliste. Eine klare, strukturierte und transparente Darstellung hilft bei der Beurteilung des Antrags und stellt einen Pluspunkt bei der Bewertung des Antrags dar. Findet in einem Antragsverfahren eine Begehung durch eine Gutachterkommission statt, wird das Thema Forschungsdaten mit hoher Wahrscheinlichkeit ebenfalls thematisiert.

Sie können auch einen Datenmanagementplan mit dem Tool Research Data Management Organizer (RDMO) der UHH <https://www.fdm.uni-hamburg.de/service/rdm.html> erstellen und Ihrem Antrag beifügen.

Fragen des Datenschutzes und der Ethik sollten Sie unbedingt vor einer Beantragung mit den zuständigen Stellen abklären.

Planen Sie in Ihrem Projekt die Entwicklung einer Datenbank und/oder Webanwendung, die als Ergebnis des Projekts publiziert werden soll? Die DFG erwartet für eine positive Beurteilung derartiger Vorhaben belastbare Aussagen zum langfristigen Betrieb über die Projektlaufzeit hinaus. Ein dauerhafter Betrieb einer solchen Anwendung erfordert laufende Anpassungen der Software und ist ein komplexer Vorgang, der ohne Personalressourcen nicht geleistet werden kann. Lassen Sie sich unbedingt vorher von uns beraten!

1. Datenbeschreibung

Auf welche Weise entstehen in Ihrem Projekt neue Daten?

Hier sind Messungen, Befragungen, Fotos, Filme, Analysen, Berechnungen usw. gemeint. Beschreiben Sie möglichst umfassend, welche Methoden Sie benutzen und welche Art von Daten dabei entstehen bzw. welche Sie nutzen werden.

Werden existierende Daten wiederverwendet?

Falls Daten verwendet werden, die nicht im Projekt entstehen, geben Sie den Kontext an, aus dem die Daten stammen, möglichst als zitierfähige Quelle (z.B. über eine DOI).

Ansonsten empfiehlt sich ein Hinweis, dass nach eigener Recherche keine geeigneten Daten zur Verfügung stehen.

Die Nutzung existierender Daten, auch wenn diese zusätzlich erfolgt oder als Vergleichsmaterial dienen, stellt in jedem Fall einen positiven Aspekt für Ihren Antrag dar.

Welche Datentypen, im Sinne von Datenformaten (z. B. Bilddaten, Textdaten oder Messdaten) entstehen in Ihrem Projekt und auf welche Weise werden sie weiterverarbeitet?

Beschreiben Sie den gesamten Ablauf der Datenerhebung, vom Datenursprung, z.B. der Messung, bis zur Nutzung. Es sollte deutlich werden, dass ein planvoller Umgang mit den Daten erfolgt. Beschreiben Sie auch, wie Sie die einzelnen Schritte dokumentieren. Hilfreich kann auch die Beschreibung einer Systematik für die Benennung von Dateiernamen und Dateien sein.

In welchem Umfang fallen diese an bzw. welches Datenvolumen ist zu erwarten?

Machen Sie einerseits Angaben zur Anzahl (z.B. der beabsichtigten Messungen, der zu führenden Interviews oder der auszuwertenden Bilder) und andererseits zur erwartbaren Spei-

chermenge pro Messung. Also z.B.: „Es sind 10 Messungen geplant, bei jeder Messung fallen ca. 10 Mbyte Daten an, für das gesamte Projekt ergibt sich ein Datenvolumen von ca. 100 Mbyte.“

Geben Sie aber auch an, welche Datenvolumina davon nur kurzfristig für die Arbeit benötigt werden und welche langfristig gespeichert werden müssen.

2. Dokumentation und Datenqualität

Welche Ansätze werden verfolgt, um die Daten nachvollziehbar zu beschreiben (z. B. Nutzung vorhandener Metadaten- bzw. Dokumentationsstandards oder Ontologien)?

Zur Datenbeschreibung gehören in jedem Fall auch Angaben über den Ursprung der Daten, verwendete Messgeräte, Tag, Uhrzeit, Ort, durchführende Personen usw. Diese Angaben sollten in einem Mess- oder Datenprotokoll dokumentiert werden.

Beschreiben Sie die Daten, geben Sie die verwendeten Einheiten an. Jemand, der Ihre Daten nicht kennt, muss anhand der vorliegenden Beschreibung die Daten verstehen und nutzen können. Verwenden Sie nur allgemein bekannte Abkürzungen oder erstellen Sie ein Abkürzungsverzeichnis.

Welche Maßnahmen werden getroffen, um eine hohe Qualität der Daten zu gewährleisten?

Machen Sie Angaben zur Eichung von Messgeräten, zu Test- und Vergleichsmessungen, Plausibilitäten usw.

Sind Qualitätskontrollen vorgesehen und wenn ja, auf welche Weise?

Beschreiben Sie die Mechanismen, mit denen Sie eine hohe Qualität der Daten sicherstellen. Es sollen Zufallsdaten oder Fehlinterpretationen verhindert werden.

Welche digitalen Methoden und Werkzeuge (z. B. Software) sind zur Nutzung der Daten erforderlich?

Dokumentieren Sie, welche Software für die Auswertung verwendet wurde. Bei Standardsoftware reicht die Angabe des Namens und der Version. Nutzen Sie selbst erstellte Software zur Auswertung, sollte diese mit entsprechender Dokumentation ggf. am Projektende zusammen mit den Daten zur weiteren Nutzung bereitgestellt werden.

Messdaten, die nur in einem spezifischen Format des Geräteherstellers vorliegen und auch nur mit diesem Gerät wieder gelesen und bearbeitet werden können, sollten entsprechend beschrieben werden. Prüfen Sie die Möglichkeit solche Daten auch in nicht proprietäre Formate zu konvertieren. So können z.B. Digitalbilder aus Kameras im herstellerspezifischen RAW-Format immer auch als DNG, TIFF oder JPEG gespeichert werden. Dies erhöht die Nachnutzbarkeit!

3. Speicherung und technische Sicherung während des Projektverlaufs

Auf welche Weise werden die Daten während der Projektlaufzeit gespeichert und gesichert?

Beschreiben Sie die genutzten Speichermedien und Abläufe. In jedem Fall müssen Backups vorgesehen werden. Wenn Sie die zentralen Netzlaufwerke des Regionalen Rechenzentrums der UHH (RRZ, www.rrz.uni-hamburg.de) verwenden, werden Backups automatisch erstellt. Bei mobilen Datenspeichern (USB-Festplatten, integrierte Speicher in Geräten, Speicherkarten, Festplatten im Laptop oder PC) müssen Sie möglichst oft eine Sicherheitskopie anfertigen. Dazu sollten Sie die mobilen Datenspeicher, die als Backupmedium genutzt werden, jeweils nur für die Erstellung des Backups mit dem Messgerät oder Laptop verbinden und diese Datenspeicher auch gesichert an einem möglichst getrennten Ort aufbewahren.

Befinden Sie sich auf einer Forschungsreise, sollten Sie immer ein tägliches Backup ihrer Daten erstellen und separat aufbewahren, vor Ort z.B. in einem Hotelsafe.

Verwenden Sie möglichst Speichermedien unterschiedlicher Hersteller und Produktionschargen, um vor möglichen Produktions- und Serienfehlern bei den Speichermedien geschützt zu sein.

Wie wird die Sicherheit sensibler Daten während der Projektlaufzeit gewährleistet (Zugriffs- und Nutzungsverwaltung)?

Beschreiben Sie, welche Personen in welcher Form Zugriff auf die Daten erhalten und welche Maßnahmen Sie treffen, damit die Daten nicht in unberechtigte Hände geraten. Speichern Sie z.B. sensible Daten auf einer USB-Festplatte, dann sollte diese Festplatte ebenfalls geschützt sein (Kennwort, ggf. Verschlüsselung), damit im Falle des Verlusts oder bei Diebstahl eine unberechtigte Nutzung der Daten ausgeschlossen ist.

Bei sensiblen Daten, insbesondere bei der Nutzung personenbezogener Daten, ist auch die Einbeziehung des Datenschutzes (<https://www.uni-hamburg.de/uhh/organisation/beauftragte/datenschutz.html>) ratsam.

4. Rechtliche Verpflichtungen und Rahmenbedingungen

Welche rechtlichen Besonderheiten bestehen im Zusammenhang mit dem Umgang mit Forschungsdaten in Ihrem Projekt?

Rechtliche Besonderheiten können bestehen bei

- personenbezogenen Daten,
- Geschäftsdaten,
- Daten mit Urheberrechtsschutz,
- Patenten,
- Daten, deren Publikation für Personen, Einrichtungen, Gruppen oder Objekte eine Gefährdung darstellen können (z.B. genaue Standortangaben bei gefährdeten Tier- und Pflanzenarten oder wertvollen Objekten; Nennung von Personen oder Gruppen bei religiöser oder politischer Gefährdung),

- Daten, die Ihnen für Ihre Forschung nur unter einschränkenden Bedingungen bereitgestellt wurden.

Beachten Sie: In illegaler Form erworbene Daten (z.B. Datenklau, Nutzung von Daten ohne Zustimmung, Fotos ohne Erlaubnis) lassen sich nicht nachträglich legalisieren und ermöglichen keine legale Forschungsarbeit. Insbesondere sind Ihre Forschungsergebnisse nicht überprüfbar.

Sind Auswirkungen oder Einschränkungen in Bezug auf die spätere Veröffentlichung bzw. Zugänglichkeit zu erwarten?

In Abhängigkeit von den zuvor genannten Rahmenbedingungen ergeben sich evtl. Einschränkungen. Beschreiben Sie diese möglichst transparent.

Auf welche Weise werden nutzungs- und urheberrechtliche Aspekte sowie Eigentumsfragen berücksichtigt?

Auch dies hängt an den zuvor genannten Rahmenbedingungen. Bei vollständig eigenen Daten können Sie eine möglichst offene Lizenz für die weitere Verwendung wählen. Nutzen Sie fremde Daten, sollten Sie, soweit möglich, auf die Quelle der Daten verweisen. Sind diese Daten z.B. schon in einem zugänglichen Repositorium gespeichert, geben Sie die URL oder DOI an, eine erneute Speicherung ist nicht notwendig.

Daten, die dem Urheberrechtsschutz unterliegen, dürfen Sie nicht einfach publizieren.

Existieren wichtige wissenschaftliche Kodizes bzw. fachliche Normen, die Berücksichtigung finden sollten?

Viele Fachwissenschaften, Forschungseinrichtungen und auch größere Projekte verfügen über Vorgaben für den Umgang mit Daten. Hier können Sie auf derartige Vorgaben und vorhandene Dokumente verweisen.

5. Datenaustausch und dauerhafte Zugänglichkeit der Daten

Welche Daten bieten sich für die Nachnutzung in anderen Kontexten besonders an?

Gut aufbereitete und dokumentierte Daten können für die weitere Forschung und die Überprüfung Ihrer Forschungsergebnisse genutzt werden. Zwischenergebnisse, Ergebnisse aus Testrechnungen oder einfach reproduzierbare Daten, z.B. aus Modellrechnungen, sollten nicht für eine Nachnutzung vorgesehen werden.

Beachten Sie die FAIR-Data Prinzipien:

https://www.forschungsdaten.org/index.php/FAIR_data_principles

Beschreiben Sie, wie Sie in dem Projekt diese Prinzipien umsetzen bzw. erfüllen werden.

Nicht alle Anforderungen müssen erfüllt sein und FAIR bedeutet keinesfalls, dass alle Daten frei verfügbar sein müssen.

Nach welchen Kriterien werden Forschungsdaten ausgewählt, um diese für die Nachnutzung durch andere zur Verfügung zu stellen?

Beschreiben Sie, warum Daten von Ihnen ausgewählt werden. So kann es z.B. sinnvoll sein die Rohdaten und die bearbeitete Endversion Ihrer Daten bereitzustellen.

Planen Sie die Archivierung Ihrer Daten in einer geeigneten Infrastruktur? Falls ja, wie und wo?

Prüfen Sie, ob es für Ihren Bereich eine sichere, auf Dauer angelegte und akzeptierte Speichermöglichkeit gibt. (z.B. GESIS in den Sozialwissenschaften, DKRZ für Klimadaten usw.). Gibt es ein solches Angebot und das Repositorium nimmt Ihre Daten an, dann sollten Ihre Daten in jedem Fall dort gespeichert werden.

Gibt es kein derartiges Angebot oder sind Ihre Daten nicht für fachwissenschaftliche Repositorien geeignet, dann nutzen Sie das Forschungsdatenrepositorium (FDR) der UHH (www.fdr.uni-hamburg.de).

Beschreiben Sie im Antrag, wo Sie welche Daten speichern werden.

Sollten Sie gar keine Archivierung und Bereitstellung vorsehen, so ist dies ausführlich zu begründen und wird in jedem Fall seitens der Gutachtenden und der DFG als kritischer Punkt angesehen werden.

Eine Ausnahme stellt nur die reine Nutzung bereits an anderer Stelle gespeicherter und verfügbarer Daten dar, was sich aber gut begründen lässt.

Gibt es Sperrfristen? Wann sind die Forschungsdaten für Dritte nutzbar?

In vielen Repositorien, so auch im FDR der UHH, können Sie Sperr- oder Embargofristen einstellen. Ihre Daten müssen nicht sofort nach Projektende verfügbar sein, Sie können auch zunächst weitere eigene Forschung oder die Erscheinungstermine relevanter Publikationen abwarten.

Im Vordergrund stehen zunächst eine potentielle Verfügbarkeit und Auffindbarkeit der Daten. Stellen Sie Ihre Forschungsdaten so offen wie möglich und so geschützt wie unbedingt notwendig bereit. Das Argument „das sind meine Daten“ ohne nachvollziehbare Gründe wird zu einer schlechteren Bewertung Ihres Antrags führen.

6. Verantwortlichkeiten und Ressourcen

Wer ist verantwortlich für den adäquaten Umgang mit den Forschungsdaten (Beschreibung der Rollen und Verantwortlichkeiten innerhalb des Projekts)?

Zunächst hat natürlich jede einzelne Forscherin und jeder Forscher die Verantwortung für ihre/seine Daten. Bei größeren Projekten bietet sich aber auch eine zentrale Verantwortung an, die darauf achtet, dass die Regelungen für den Umgang mit Daten eingehalten werden. In einem Antrag ist es ein positives Signal, wenn derartige Regelungen und Verantwortlichkeiten vorgesehen werden.

Auch der Hinweis auf eine enge Kooperation mit dem FDM-Zentrum (evtl. auch bestätigt durch einen Letter of Intent (LOI)) untermauert das Vorgehen.

In Sonderforschungsbereichen, Forschungsgruppen und Graduiertenkollegs sollten auch immer Ressourcen für die Schulung der Graduiierenden vorgesehen werden. Auch hier kann das FDM-Zentrum Sie unterstützen.

Welche Ressourcen (Kosten, Zeit oder anderes) sind erforderlich, um einen adäquaten Umgang mit Forschungsdaten im Projekt umzusetzen?

Nutzen Sie die bestehenden Fördermöglichkeiten der DFG, insbesondere in großen Verbundprojekten bietet sich auch die Schaffung einer Stelle für den Bereich IT und Datenmanagement an. Betrachten Sie Forschungsdatenmanagement nicht als Nebensache, die „irgendwie“ erledigt wird. Die DFG und die Gutachtenden messen diesem Thema eine hohe Bedeutung bei.

Fallen Kosten für die Speicherung der Daten an? Normalerweise stehen Ressourcen über das RRZ kostenfrei zur Verfügung. Besonders große Bedarfe können aber durchaus auch eine finanzielle Beteiligung erfordern, dies sollte vorher mit dem RRZ abgesprochen werden.

Die langfristige Speicherung der Daten im UHH-FDR ist kostenfrei. Erfolgt die Speicherung von Daten bei Drittanbietern, können aber evtl. Kosten anfallen. Insbesondere sind hier dauerhafte Kosten nach der Projektlaufzeit vorher abzuklären, da diese in der Regel nicht von der UHH übernommen werden. Auch dauerhafte Betriebskosten und Aufwände für die Pflege von Webseiten und -anwendungen und Datenbanken werden nicht von der UHH übernommen.

Wer ist nach Ende der Laufzeit des Projekts für das Kuratieren der Daten verantwortlich?

Mit einer Kuration der Daten ist die weitere inhaltliche Pflege der Daten gemeint, z.B. die Ergänzung oder Korrektur von Daten in einer Datenbank, aber auch die dauerhafte Nutz- und Interpretierbarkeit, die z.B. durch die Migration in andere Datenformate erreicht werden kann.

In der Regel werden nach der Projektlaufzeit keine oder nur geringe Ressourcen für diese Aufgaben mehr vorhanden sein. Zu unterscheiden sind zwei grundlegende Fälle:

- **Datenbanken und Webanwendungen**
Nach einer entsprechenden gemeinsamen Planung am Projektanfang und Übergabe einer Dokumentation und der notwendigen Zugänge am Projektende an das FDM-Zentrum oder bei der Nutzung der vom FDM-Zentrum angebotenen Heurist-Datenbankinstanz, wird der Betrieb der Anwendung, soweit technisch möglich, gewährleistet. Die Inhalte der Datenbank oder der Webanwendung können von Ihnen weiter gepflegt werden. Die Kuration der Inhalte wird aber nicht vom FDM-Zentrum übernommen.
- **Forschungsdaten, die im FDR der UHH oder einem anderen Repository gespeichert werden**
Die Betreiber von Repositorien verändern die gespeicherten Daten in der Regel nicht. Es findet keine Migration in neue Dateiformate statt, vielmehr wird der physische Erhalt, die sog. bitstream preservation Methode angewandt. Die dauerhafte Nutzbarkeit und Interpretierbarkeit der Daten kann somit nicht garantiert werden, da diese von der Verfügbarkeit der jeweiligen Datenformate bzw. der notwendigen Software abhängen kann. Eine Kuration der Inhalte, also z.B. eine Fehlerkorrektur oder Ergänzung, kann nicht erfolgen. Korrigierte und geänderte Daten können nur als neue Version der Daten gespeichert werden. Dieses Vorgehen entspricht auch den Regularien der DOI-Vergabe.